Psychometric Properties of the Indonesian Voice Handicap Index-10 among Elementary School Teachers in Bandar Lampung City, Indonesia

Fivien Fedriani[1*], Fitria Saftarina[2], Reni Zuraida[3], Dyah Wulan Sumekar Rengganis Wardani[4], Betta Kurniawan[5]
[1,2,3,4,5]Universitas Lampung, Bandar Lampung, Indonesia
Email: fivienfedriani@gmail.com

Abstract

Teachers frequently experience voice-related difficulties because their profession entails sustained vocal use, making a brief yet psychometrically sound screening instrument essential in school settings. Although the Indonesian adaptation of the Voice Handicap Index-10 (VHI-10) has undergone evaluation in ear, nose, and throat outpatient populations, its measurement performance in the teaching workforce has yet to be established. This study assessed the extent to which the Indonesian VHI-10 captures perceived voice-related limitations among primary school educators in Bandar Lampung City. Data were gathered in September 2026 from 305 teachers representing 14 elementary schools across seven sub-districts. Internal consistency was assessed using Cronbach's alpha, ordinal alpha, and McDonald's omega, while the instrument's underlying structure was explored through parallel analysis and ordinal single-factor confirmatory factor analysis (CFA). Construct validity was further appraised by examining predefined differences between groups. Participants obtained an average VHI-10 score of 3.9 ± 4.8, with 34.8% recording no perceived voice handicap. Reliability estimates indicated satisfactory consistency, yielding Cronbach's alpha of 0.85 (95% CI 0.83–0.88), ordinal alpha of 0.95, and omega of 0.96; corrected item–total coefficients ranged from 0.42 to 0.74. A single latent dimension accounted for 68.0% of the total variance, while CFA demonstrated an adequate model fit (CFI = 0.985, TLI = 0.980, RMSEA = 0.067, SRMR = 0.069), with factor loadings spanning 0.69–0.93. Consistent with the anticipated relationships, higher scores were observed among educators exposed to frequent shouting, overcrowded classrooms, excessive ambient noise, and associated health complaints. Scores also escalated alongside the cumulative burden of occupational exposures, whereas no meaningful variation emerged across sex or age categories. Overall, the Indonesian VHI-10 demonstrates strong reliability and a coherent single-factor structure, supporting its application in identifying voice-related difficulties among elementary school teachers. Nevertheless, the substantial concentration of zero scores warrants attention when interpreting results at the lower end of the measurement scale.

Keywords: *confirmatory factor analysis; reliability; teachers; validity; Voice Handicap Index-10.*

– – – – – – – – – – ◆ – – – – – – – – – –

A. INTRODUCTION

Teaching relies heavily on continuous vocal activity, making educators particularly susceptible to voice-related disorders, especially dysphonia. A global meta-analysis reported that approximately 37.7% of schoolteachers experience voice disorders (Baghban et al., 2025). Meanwhile, reported proportions reached 19.4% in Madinah (Almuzaini, 2025) and 21% in northern Ethiopia (Belete et al., 2025). Evidence from a nationwide Taiwanese cohort likewise indicated that schoolteachers

face a greater incidence of vocal disturbances than university lecturers (Chen et al., 2022). An additional concern is that many educators experiencing vocal difficulties do not seek professional medical assistance (Almuzaini, 2025). This situation highlights the importance of having a concise, independently completed screening instrument with established measurement accuracy for use in educational institutions and workplace health programs.

The original Voice Handicap Index (VHI) comprises 30 statements designed to capture the consequences of voice impairment across three dimensions: functional limitations, physical manifestations, and emotional distress (Jacobson et al., 1997). A condensed alternative, known as the VHI-10, consists of ten selected statements demonstrating stronger measurement performance and produces an overall score ranging from 0 to 40 (Rosen et al., 2004). Among individuals without vocal complaints, the reported reference average is 2.83 (Arffa et al., 2012). Adaptations of this instrument have subsequently appeared in various languages. The Brazilian Portuguese edition demonstrated a Cronbach's alpha of 0.87 (Costa et al., 2013), whereas the Turkish adaptation established threshold values for screening purposes (Tahir et al., 2025). Its Malay counterpart has also been applied to determine the occurrence of voice disorders among secondary-level educators (Moy et al., 2015).

Within Indonesia, an adapted VHI-10 underwent validation involving patients attending otolaryngology outpatient services, producing Pearson item-correlation coefficients between 0.50 and 0.85 and an internal consistency coefficient of 0.78 (Fedriani et al., 2026a). Nevertheless, an instrument's measurement characteristics cannot automatically be assumed to remain equivalent when transferred to a different population (Mokkink et al., 2010; Prinsen et al., 2018). Unlike clinical patients, teachers represent a working population in which vocal complaints may be relatively mild or absent, potentially resulting in an accumulation of minimum scores and asymmetrical response patterns. Such characteristics may influence how individual questionnaire items perform. Furthermore, the underlying factor structure of the Indonesian VHI-10 has not previously been scrutinized through confirmatory factor analysis. To address these limitations, the present investigation evaluated the distribution of questionnaire scores, internal consistency, dimensional structure, and known-groups validity of the Indonesian VHI-10 among elementary school educators in Bandar Lampung City.

B. METHOD

This methodological investigation employed a cross-sectional psychometric design and took place in September 2026 across 14 government-operated and privately managed elementary schools in Bandar Lampung, Indonesia. The dataset originated from a broader investigation concerning dysphonia in elementary school educators, while the associated factors underlying this condition in the same population are discussed in a separate publication (Fedriani et al., 2026b). Participant recruitment followed three successive non-probability selection

procedures. Initially, seven out of 20 sub-districts were deliberately chosen based on their proximity to heavily trafficked arterial roads and the availability of an adequate number of schools and teaching personnel. Subsequently, participating schools situated along these major routes were identified within each selected area. Finally, every currently employed teacher at the participating institutions was invited to take part.

Participation was restricted to educators actively engaged in classroom instruction, capable of reading Indonesian and prepared to provide written consent. Individuals were not included if they had experienced speech or voice abnormalities before entering the teaching profession, undergone thyroid or vocal fold surgery during the preceding six months, had an autoimmune condition, or were experiencing an acute respiratory infection during fieldwork. The final dataset comprised 305 teachers, surpassing the minimum sample of 100 recommended for examining internal consistency and factorial configuration (Terwee et al., 2007).

The Indonesian adaptation of the VHI-10 contains ten items addressing the consequences of vocal difficulties. Each statement is answered using a five-level response format, ranging from 0 (never) to 4 (always), producing an aggregate score between 0 and 40; higher values denote more pronounced voice-related impairment (Fedriani et al., 2026a; Rosen et al., 2004). Respondents filled out the instrument independently. An additional structured form gathered demographic and occupational information, alongside variables intended for known-groups assessment. These included routinely raising one's voice or shouting during lessons, classroom enrolment exceeding 28 pupils versus 28 or fewer in accordance with the nationally stipulated class-size ceiling (Kementerian Pendidikan Dasar dan Menengah, 2026), perceived levels of classroom disturbance, sustained vocal activity without resting intervals (never, sometimes, or often), and previous occurrences of upper respiratory tract infections, allergic rhinitis, and gastroesophageal or laryngopharyngeal reflux. Ambient classroom sound was recorded during teaching sessions through the NIOSH Sound Level Meter smartphone application.

Score concentration at either extreme was examined through floor and ceiling effects, which were deemed substantial when over 15% of respondents obtained the minimum or maximum attainable value, respectively (Terwee et al., 2007). Measurement consistency was examined through three complementary coefficients: Cronbach's alpha accompanied by its 95% confidence interval (CI), ordinal alpha derived from polychoric associations (Zumbo et al., 2007), and McDonald's omega (McNeish, 2018). Coefficients reaching 0.70 or above were interpreted as satisfactory (Terwee et al., 2007). Individual statements were considered sufficiently associated with the overall construct when their corrected item–total correlations reached at least 0.30. Prior to extracting the latent structure, sampling adequacy was inspected using the Kaiser–Meyer–Olkin (KMO) measure alongside Bartlett's test of sphericity. The appropriate number of dimensions was subsequently determined through parallel analysis based on polychoric

correlations.

The hypothesized single-dimensional configuration was subsequently examined through confirmatory factor analysis (CFA), employing the weighted least squares mean- and variance-adjusted (WLSMV) estimation procedure, an approach suited to ordinal response formats (Li, 2016). Since the highest response options (3 and 4) occurred infrequently in several statements, these categories were consolidated with category 2, resulting in three ordered response levels: 0, 1, and ≥ 2 for ordinal modelling. Model adequacy was judged using the comparative fit index (CFI) and Tucker–Lewis index (TLI), with values of at least 0.95 denoting good correspondence, alongside the root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR), for which values not exceeding 0.08 indicated acceptable correspondence (Hu & Bentler, 1999). Robust fit statistics appropriate for categorical observations were also presented (Savalei, 2021). Evidence of convergent validity was assessed through average variance extracted (AVE) of at least 0.50 and composite reliability (CR) of at least 0.70 (Fornell & Larcker, 1981).

For the assessment of known-groups validity, the initial hypotheses anticipated greater VHI-10 scores among educators encountering recognized occupational or health-related contributors to voice disorders (Byeon, 2019; Jiang et al., 2024). A further expectation was that questionnaire scores would progressively rise as exposure to occupational hazards accumulated, particularly habitual shouting, classrooms containing more than 28 pupils, and perceived excessive classroom noise. Comparisons involving two independent groups were conducted using the Mann–Whitney U procedure, with effect magnitude expressed as $r = Z/\sqrt{N}$. Differences involving multiple groups were evaluated through the Kruskal–Wallis test. Statistical significance was established using a two-tailed threshold of 0.05. All computations were carried out in R version 4.3.3 with the psych and lavaan packages (Rosseel, 2012).

Approval to conduct the investigation was secured from the Research Ethics Committee of the Faculty of Medicine, Universitas Lampung (No. 5315/UN26.18/PP.05.02.00/2026). Authorization from the Bandar Lampung City Education and Culture Office was also obtained (No. 000.9.2/2751/III.01/2026). Written informed consent was collected from every participating teacher before any research activities commenced.

C. RESULT AND DISCUSSION

1. Participant Characteristics and Score Distribution

Complete responses were obtained from all 305 participating teachers, leaving no unanswered VHI-10 items. Participants had an average age of 41.1 ± 9.9 years, spanning 21 to 60 years. Women constituted the larger proportion of the sample (83.3%), while 75.4% were assigned as classroom teachers (Table 1).

Table 1 Characteristics of participants and distribution of VHI-10 scores (n = 305)

Characteristic	n (%)
Age, years, mean ± SD (range)	41.1 ± 9.9 (21–60)
Female	254 (83.3)
Bachelor's degree	294 (96.4)
Class teacher	230 (75.4)
Habitual loud speaking/shouting	197 (64.6)
> 28 students per class	121 (39.7)
Perceived noisy classroom	148 (48.5)
VHI-10 total score, mean ± SD	3.9 ± 4.8
VHI-10 total score, median (IQR); range	2 (0–6); 0–26
Score of 0 (floor)	106 (34.8)
Score of 40 (ceiling)	0 (0.0)
Score ≥ 7	69 (22.6)

Source: Primary data (2026). SD = standard deviation; IQR = interquartile range

The overall score pattern showed a marked right-sided skew (skewness = 1.52). A pronounced floor concentration was observed, with 34.8% of respondents obtaining a zero score, while none attained the highest possible value. Item-level responses were most common for item 2, concerning difficulty being understood amid background noise (51.5%), followed by item 7, describing inconsistent vocal clarity (37.0%), and item 1, relating to difficulty making one's voice heard (34.4%). Conversely, item 9, which addresses the feeling of being limited by one's voice, received the fewest affirmative responses (6.9%) (Table 2).

Table 2 Item characteristics, item–total correlations and factor loadings of the Indonesian VHI-10 (n = 305)

Item	Mean ± SD	Score ≥ 1 (%)	CITC	*α* if item deleted	Loading
1. "My voice makes it difficult for people to hear me"	0.53 ± 0.85	34.4	0.50	0.846	0.69
2. "People have difficulty understanding me in a noisy room"	0.88 ± 0.97	51.5	0.54	0.846	0.77
3. "My voice difficulties restrict my personal and social life"	0.40 ± 0.78	25.6	0.74	0.822	0.91
4. "I feel left out of conversations because of my voice"	0.21 ± 0.57	15.7	0.63	0.837	0.85
5. "My voice problem causes me to lose income"	0.13 ± 0.45	10.2	0.52	0.846	0.81
6. "I feel as though I have to strain to produce voice"	0.53 ± 0.96	29.5	0.67	0.829	0.84
7. "The clarity of my voice is unpredictable"	0.62 ± 0.93	37.0	0.60	0.837	0.84
8. "My voice problem upsets me"	0.29 ± 0.67	19.7	0.65	0.832	0.87
9. "My voice makes me feel handicapped"	0.08 ± 0.32	6.9	0.55	0.849	0.93
10. "People ask, 'What's wrong with your voice?'"	0.21 ± 0.51	16.4	0.42	0.850	0.70

Source: Primary data (2026). CITC, item–rest correlation; loading, standardized coefficient of the single-factor WLSMV model; all loadings $p < 0.001$. Item wording is quoted from the original English VHI-10 (Rosen et al., 2004).

The floor effect is expected because the questionnaire was designed for a clinical population, and most teachers in a workplace survey have little or no handicap. The mean score of 3.9 was close to the normative mean of 2.83 reported among people without voice complaints (Arffa et al., 2012). The floor effect limits discrimination among teachers with minimal handicap, but it does not prevent its use for screening, the aim of which is to identify teachers with scores above a threshold. The most frequently endorsed items concerned being understood when the classroom is loud and unpredictable vocal clarity, which reflect early, functional voice complaints commonly reported by teachers (Almuzaini, 2025; Moy et al., 2015).

2. Internal Consistency

The instrument demonstrated satisfactory internal reliability, yielding a Cronbach's alpha of 0.85 (95% CI 0.83–0.88), an ordinal alpha of 0.95, and a McDonald's omega of 0.96. Corrected item–total coefficients extended from 0.42 for item 10 to 0.74 for item 3. Removing individual statements produced no improvement in Cronbach's alpha, which remained between 0.82 and 0.85. Average inter-item associations were 0.40 (range 0.16–0.63) under Pearson correlations and 0.64 (range 0.34–0.82) under polychoric correlations. Reliability estimates also remained relatively consistent across demographic subgroups, with alpha values of 0.83 among male and 0.86 among female teachers, compared with 0.85 in those aged ≤ 40 years and 0.86 in those aged > 40 years.

Cronbach's alpha of 0.85 was higher than the 0.78 reported among Indonesian otolaryngology patients (Fedriani et al., 2026a) and comparable to the 0.87 of the Brazilian Portuguese versions (Costa et al., 2013). Ordinal alpha and omega were higher still, because Cronbach's alpha computed from Pearson correlations underestimate's reliability when items are ordinal and strongly skewed, as in this non-clinical sample (McNeish, 2018; Zumbo et al., 2007). Removing any single item did not raise alpha, and every item correlated above 0.30 with the rest of the scale; therefore, none of the ten items needs to be dropped. The similar reliability across sex and age groups suggests that the instrument performs consistently in these subgroups.

3. Structural Validity

Preliminary diagnostics confirmed that the dataset was well-suited to latent-factor examination, as reflected by a KMO value of 0.87, individual sampling adequacy indices ranging from 0.81 to 0.91, and a significant Bartlett's test (χ^2 = 1,197.1, df = 45, $p < 0.001$). The polychoric correlation matrix yielded an initial eigenvalue of 6.80, accounting for 68.0% of total variance, while the subsequent eigenvalue declined to 0.93. Parallel analysis consequently supported the extraction of a single underlying dimension. Confirmatory factor analysis further indicated an acceptable one-factor configuration, with standardized factor coefficients ranging from 0.69 to 0.93 (Table 3). The most prominent modification index (27.9) concerned correlated residuals between items 1 and 2, both addressing difficulties in making one's voice audible or intelligible to other people.

Table 3 Reliability and structural validity of the Indonesian VHI-10 (n = 305)

Measure	Result	Criterion
Reliability		
Alpha (95% CI)	0.85 (0.83–0.88)	≥ 0.70
Alpha, ordinal (polychoric)	0.95	≥ 0.70
McDonald's omega	0.96	≥ 0.70
Corrected item–total correlation	0.42–0.74	≥ 0.30
Structural validity (one-factor CFA, WLSMV)		

χ^2 (df); p	82.5 (35); < 0.001	–
Comparative fit index	0.985	≥ 0.95
Tucker–Lewis index	0.980	≥ 0.95
RMSEA (90% interval)	0.067 (0.048–0.086)	≤ 0.08
Standardized RMR	0.069	≤ 0.08
Robust CFI; TLI	0.948; 0.934	≥ 0.95
Robust RMSEA (90% CI)	0.121 (0.000–0.203)	≤ 0.08
Standardized loadings	0.69–0.93	≥ 0.50
AVE	0.68	≥ 0.50
Composite reliability	0.96	≥ 0.70

Source: Primary data (2026). Abbreviations are defined in the Methods section.

Parallel analysis and CFA both pointed to one dimension, which justifies summing the ten items into a single score (Rosen et al., 2004). All loadings were high, and the AVE and composite reliability indicated that the items share a common construct of voice handicap. The highest loadings were found for item 9 (feeling handicapped) and item 3 (restriction of personal and social life), whereas the lowest were for items 1 and 10, which describe how others perceive the voice. The residual correlation between items 1 and 2 suggests shared content related to being heard and understood, which is particularly relevant for teachers working in noisy classrooms. The standard WLSMV indices fell within the cut-offs proposed by Hu and Bentler (1999), whereas the robust indices were weaker and the robust RMSEA had a wide confidence interval. Such indices are sensitive to sparse response categories and a small number of degrees of freedom (Savalei, 2021); fit should therefore be interpreted as acceptable rather than excellent.

4. Known-Groups Validity

The findings were consistent with the prespecified known-groups expectations (Table 4). Higher VHI-10 scores were observed among teachers who frequently raised their voices, handled classes exceeding 28 pupils, considered their teaching environment acoustically disruptive, or experienced upper respiratory tract infections, allergic rhinitis, or reflux. These differences were associated with small-to-moderate effect magnitudes (r = 0.12–0.25). Scores also rose in tandem with more frequent extended speaking periods without breaks and a greater accumulation of workplace-related exposures, progressing from a median of 0 among teachers reporting none to 5 among those exposed to all three. No meaningful score variation emerged across sex or age categories. Class size showed a positive association with VHI-10 scores ($\rho = 0.28$, $p < 0.001$), whereas objectively recorded classroom sound intensity (72.1 ± 6.5 dB) showed virtually no relationship ($\rho = -0.02$, $p = 0.77$).

Table 4 Known-groups validity of the Indonesian VHI-10 total score (n = 305)

Group	n	Mean ± SD	Median (IQR)	p	r
Occupational exposures					
Habitual loud speaking/shouting	197	4.7 ± 5.2	3 (0–7)	< 0.001*	0.25
No shouting	108	2.4 ± 3.6	1 (0–3.25)		
> 28 students per class	121	5.2 ± 5.4	3 (1–9)	< 0.001*	0.23
≤ 28 students per class	184	3.0 ± 4.2	1 (0–5)		
Perceived noisy classroom	148	4.9 ± 5.4	3 (0–7.25)	0.001*	0.19
Perceived quiet classroom	157	3.0 ± 4.0	2 (0–4)		
Prolonged speaking without pauses: often	22	7.4 ± 7.8	5 (1.5–9.25)	< 0.001*[a]	

Sometimes	154	4.1 ± 4.4	2.5 (0.25–6)		
Never	129	3.0 ± 4.4	1 (0–4)		
Number of occupational exposures[b]: 3	57	6.4 ± 6.1	5 (2–10)	< 0.001*[a]	
2	99	4.6 ± 5.0	3 (0–7)		
1	97	2.9 ± 3.7	2 (0–4)		
0	52	1.7 ± 3.3	0 (0–1.25)		
Health history					
Upper respiratory tract infection	64	5.7 ± 6.3	3.5 (0–10)	0.018*	0.14
None	241	3.4 ± 4.3	2 (0–5)		
Allergic rhinitis	75	5.2 ± 5.7	3 (1–8.5)	0.012*	0.14
None	230	3.5 ± 4.5	2 (0–5)		
Gastroesophageal/laryngopharyngeal reflux	72	5.1 ± 5.5	4 (0–8)	0.031*	0.12
None	233	3.5 ± 4.6	2 (0–5)		
Sociodemographic					
Female	254	3.9 ± 4.8	2 (0–6)	0.94	0.00
Male	51	3.9 ± 4.9	2 (0–5)		
Age > 40 years	149	4.1 ± 5.0	2 (0–6)	0.49	0.04
Age ≤ 40 years	156	3.7 ± 4.7	2 (0–6)		

Source: Primary data (2026). *Significant at α = 0.05 (Mann–Whitney U test unless marked otherwise); aKruskal–Wallis test; bhabitual shouting, > 28 students per class and perceived noisy classroom. R = effect size (Z/√N); IQR = interquartile range; SD = standard deviation.

These differences are presented as evidence of construct validity, that is, that the questionnaire separates teachers who, on theoretical grounds, should report more voice handicap from those who should report less, rather than as estimates of risk factors; the determinants of dysphonia in this sample, analysed with dichotomized outcomes and multivariable adjustment, are reported separately (Fedriani et al., 2026b). The direction of the differences is consistent with risk factors identified in meta-analyses (Byeon, 2019; Jiang et al., 2024) and with evidence that teachers raise their voice in noisy environments (Nusseck et al., 2022) and that classroom acoustics affect teachers' health (Mealings et al., 2024; Sharp & Cook, 2024). Allergic rhinitis was associated with higher continuous scores in the present analysis, although it was not significantly associated with dysphonia defined by a cut-off score; comparisons of continuous scores are more sensitive to modest shifts across the whole distribution than comparisons of a dichotomized outcome. The VHI-10 score was not correlated with the measured noise level, possibly because noise was measured only once and almost all classrooms exceeded the 55 dBA standard for school areas (Kementerian Lingkungan Hidup, 1996), which limited the variation in exposure, whereas perceived noise reflects teachers' cumulative daily experience.

5. Limitations

Several methodological constraints should be acknowledged. Because the questionnaire was completed on a single occasion, the study could not establish score stability over time or quantify measurement error. Criterion validity and a defensible threshold score also remained undetermined, given the absence of laryngoscopic findings, clinician-rated voice evaluations, and an additional voice measure for assessing convergent validity. Ordinal analyses further necessitated the consolidation of certain response options. Moreover, participants were recruited through non-random school selection, limited to institutions situated along major roads, and the dataset originated from a broader

investigation of dysphonia. Subsequent research would benefit from repeat administrations, direct clinical appraisal of voice function, and participant recruitment across a wider range of Indonesian regions.

D. CONCLUSION

The Indonesian VHI-10 showed adequate internal consistency, a unidimensional structure with high factor loadings and good known-groups validity among elementary school teachers in Bandar Lampung City, although a floor effect was present. The questionnaire is suitable as a quick first-step screen for voice handicap in occupational health programmes for teachers. Further studies should evaluate its test–retest reliability and determine a cut-off point against clinical voice examination in teachers.

E. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Bandar Lampung City Education and Culture Office, the headteachers of the participating primary schools, and every teacher who contributed to this research. This investigation received no financial backing from external sources, and the authors declare no competing interests.

REFERENCES

1. Almuzaini, H. I. (2025). Voice problems in school teachers: A cross-sectional study from Madinah, Saudi Arabia. Noise and Health, 27(129), 760–769. https://doi.org/10.4103/nah.nah_187_25

2. Arffa, R. E., Krishna, P., Gartner-Schmidt, J., & Rosen, C. A. (2012). Normative values for the Voice Handicap Index-10. Journal of Voice, 26(4), 462–465. https://doi.org/10.1016/j.jvoice.2011.04.006

3. Baghban, K., Golmohammadi, G., & Asadollahpour, F. (2025). The worldwide prevalence of voice disorders among schoolteachers: A systematic review and meta-analysis. Journal of Voice. Advance online publication. https://doi.org/10.1016/j.jvoice.2025.04.018

4. Belete, A. K., Yenealem, D. G., & Azanaw, J. (2025). Self-reported prevalence and associated factors of work related voice disorders among school teachers in Sekota town, Wag Himra zone, North Ethiopia, 2021: A cross-sectional survey. BMC Public Health, 25, 2639. https://doi.org/10.1186/s12889-025-23850-6

5. Byeon, H. (2019). The risk factors related to voice disorder in teachers: A systematic review and meta-analysis. International Journal of Environmental Research and Public Health, 16(19), 3675. https://doi.org/10.3390/ijerph16193675

6. Chen, B. L., Cheng, Y. Y., Lin, C. Y., & Guo, H. R. (2022). Incidence of voice disorders among private school teachers in Taiwan: A nationwide longitudinal study. International Journal of Environmental Research and Public Health, 19(3), 1130. https://doi.org/10.3390/ijerph19031130

7. Costa, T., Oliveira, G., & Behlau, M. (2013). Validation of the Voice Handicap Index: 10 (VHI-10) to the Brazilian Portuguese. CoDAS, 25(5), 482–485. https://doi.org/10.1590/S2317-17822013000500013

8. Fedriani, F., Apri Yanti, L., Ghanie, A., & Bahar, E. (2026a). Uji validitas dan reliabilitas Voice Handicap Index-10 (VHI-10) adaptasi bahasa Indonesia di Bagian T.H.T.K.L RSUP Dr. Mohammad Hoesin Palembang [Validity and reliability of the Indonesian adaptation of the Voice Handicap Index-10 (VHI-10) at the Otorhinolaryngology–Head and Neck Surgery Department, Dr. Mohammad Hoesin General Hospital, Palembang]. JK Unila (Jurnal Kedokteran Universitas Lampung), 10, 67–77. https://doi.org/10.23960/jkunila.v10i1.pp67-77

9. Fedriani, F., Saftarina, F., Zuraida, R., Wardani, D. W. S. R., & Kurniawan, B. (2026b). Class size and classroom noise as determinants of dysphonia among elementary school teachers in Indonesia: A cross-sectional study using the Indonesian Voice Handicap Index-10 [Manuscript submitted for publication].

10. Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50. https://doi.org/10.1177/002224378101800104

11. Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

12. Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The Voice Handicap Index (VHI): Development and validation. American Journal of Speech-Language Pathology, 6(3), 66–70. https://doi.org/10.1044/1058-0360.0603.66

13. Jiang, G., Xu, X., Yu, K., Tang, G., Yang, R., & Li, X. (2024). The risk factors of voice disorders in non-university teachers: A systematic review and meta-analysis. Journal of Voice. Advance online publication. https://doi.org/10.1016/j.jvoice.2024.10.023

14. Kementerian Lingkungan Hidup. (1996). Keputusan Menteri Negara Lingkungan Hidup Nomor KEP-48/MENLH/11/1996 tentang baku tingkat

kebisingan [Decree of the State Minister of Environment No. KEP-48/MENLH/11/1996 on noise level standards]. https://jdih.menlhk.go.id/

15. Kementerian Pendidikan Dasar dan Menengah. (2026). Keputusan Menteri Pendidikan Dasar dan Menengah Republik Indonesia Nomor 14 Tahun 2026 [Decree of the Minister of Primary and Secondary Education No. 14 of 2026 on the number of students per study group]. Kementerian Pendidikan Dasar dan Menengah.

16. Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. Behavior Research Methods, 48(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

17. McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. Psychological Methods, 23(3), 412–433. https://doi.org/10.1037/met0000144

18. Mealings, K., Maggs, L., & Buchholz, J. M. (2024). The effects of classroom acoustic conditions on teachers' health and well-being: A scoping review. Journal of Speech, Language, and Hearing Research, 67(1), 346–367. https://doi.org/10.1044/2023_JSLHR-23-00256

19. Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. Journal of Clinical Epidemiology, 63(7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006

20. Moy, F. M., Hoe, V. C. W., Hairi, N. N., Chu, A. H. Y., Bulgiba, A., & Koh, D. (2015). Determinants and effects of voice disorders among secondary school teachers in Peninsular Malaysia using a validated Malay version of VHI-10. PLoS ONE, 10(11), e0141963. https://doi.org/10.1371/journal.pone.0141963

21. Nusseck, M., Immerz, A., Richter, B., & Traser, L. (2022). Vocal behavior of teachers reading with raised voice in a noisy environment. International Journal of Environmental Research and Public Health, 19(15), 8929. https://doi.org/10.3390/ijerph19158929

22. Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of Life Research, 27(5), 1147–1157. https://doi.org/10.1007/s11136-018-1798-3

23. Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., & Murry, T. (2004). Development and validation of the Voice Handicap Index-10. The Laryngoscope, 114(9), 1549–1556. https://doi.org/10.1097/00005537-200409000-00009

24. Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. Journal of Statistical Software, 48(2), 1–36. https://doi.org/10.18637/jss.v048.i02

25. Savalei, V. (2021). Improving fit indices in structural equation modeling with categorical data. Multivariate Behavioral Research, 56(3), 390–407. https://doi.org/10.1080/00273171.2020.1717922

26. Sharp, E., & Cook, R. (2024). Voice symptoms and wellbeing in school teachers in England. Journal of Voice, 38(5), 1252.e1–1252.e10. https://doi.org/10.1016/j.jvoice.2022.02.005

27. Tahir, E., Yaşar, Ö., Erensoy, İ., & Kemal, Ö. (2025). Voice Handicap Index-10 and voice-related quality of life: Determination of cutoff points in Turkish language for screening of voice disorders. Journal of Voice. Advance online publication. https://doi.org/10.1016/j.jvoice.2025.03.019

28. Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology, 60(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

29. Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. Journal of Modern Applied Statistical Methods, 6(1), 21–29. https://doi.org/10.22237/jmasm/1177992180